

比较在时间缩放和俯仰平移算法中海明窗函数和布莱克曼窗函数的效果

林志伟¹, 达力¹, 王浩², 韩炜², 林凡²

(1. 厦门大学 计算机科学系 福建 厦门 361005 2. 厦门大学 软件学院 福建 厦门 361005)

【摘要】: 实时变调过程被广泛用于各类音乐制作。变调技术可分为两大类: 时间域类型和频域类型。在大规模变换计算方面频域法与时域方法相比更具有优势。然而, 在频域处理使用傅里叶变换, 导致不可避免的频率泄漏会影响音高转移效果的准确性。为了抑制傅立叶变换的副作用, 窗函数被用于减低频谱混叠。在实际处理中, 海明窗函数和布莱克曼窗函数经常使用。在本文中, 我们比较了两个窗口函数在频率泄露的抑制方面的性能和精度。

【关键词】: 俯仰平移算法、海明窗函数、布莱克曼窗函数

1. 介绍

从频率的角度, 音频可以被看作是由不断改变的正弦波组成的离散信号。音乐信号可以看成是一个短期(一般为 10~30 毫秒)的平滑信号。它在一段时间内的相对稳定和简单的。同时这种语音在主观上是单调的声音。由于音乐的这种稳定特性, 短时傅立叶变换(STFT)^[2]被广泛使用。这种信号通常被称为一段时期内的一帧。该信号采用窗口移动方法, 可以拦截所有的时间帧。

2. 时间/频率改变算法

改变音频的音调, 就是改变包含了音频波的频率。音调转换算法正是基于这一思想的。双波频, 它在理论上将增加音调音乐的 G-8 度。然而, 纯间距在频域的改变, 使每个阶段的不一致。并且它会产生回声效果。此外, 在理论上, 因为有限的窗口长度, 频率不能被严格区分开来, 即频谱走样。这个理由使我们不能在波谱中准确地分析波的组成。频率泄漏和频谱走样是一个问题的两个报表。本文不区分上述两种报表。

短时傅里叶变换(SFFT)分析综合法是一个有效解决相位不连续的解决方案。这种方法利用了窗口增量, 傅里叶变换, 频率/相匹配, 综合了窗口和堆放的过程^[3]。有效地消除了回声效果, 这被称为相合成。

2.1 改进相合成算法

传统的相位模型包括四个过程。有窗口信号傅立叶变换, 频率和相位相匹配, 综合窗口和输出叠加^[4]。在本文中, 我们使用海明窗作为一个离散傅立叶变换操作。之后比较不同的音频样本。发现效果优于布莱克曼窗。实验还表明, 海明窗在抑制频率泄漏有较好的效果。序列在窗函数后需要重建, 以恢复其原有的能源。Blackman 窗和海明窗, 重建的过程中可以使用相同的窗函数。通过使用综合叠加的方法, 它可以恢复时域信号。

布莱克曼窗

$$w_h(n) = \begin{cases} 0.42 - 0.5 \cos \frac{2\pi n}{M} + 0.08 \cos \frac{4\pi n}{M} \\ 0 \end{cases} \quad (1)$$

海明窗

$$w_m(n) = \begin{cases} 0.54 - 0.46 \cos \frac{2\pi n}{M} & 0 \leq n < M \\ 0 & \text{其它} \end{cases} \quad (2)$$

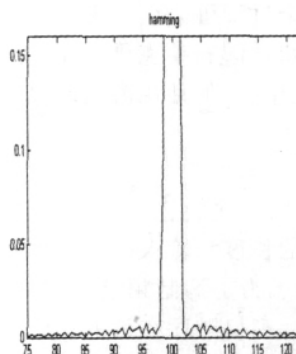


图 1. 使用海明窗和傅里叶转换的音频信号频谱

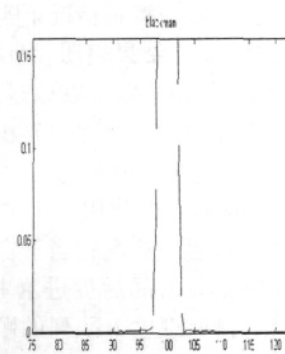


图 2. 使用布莱克曼窗和傅里叶转换的音频信号频谱

通过比较, 海明窗将信号频率限制在一个的接近 100 赫兹的很窄的区域。信号采用海明窗口呈现出更窄的主要豆、能量集中和更准确的频率等特征, 这将提高性能变调过程功能。布莱克曼窗显然在能量泄漏方面有更广泛的主豆。同时, 应该指出采用海明窗的信号在实验中有更多的旁瓣, 这在一定程度上抵消了狭窄主豆的浓度效应。通常, 在频率泄漏的约束方面海明窗优于布莱克曼窗。

2.2 时域窗口及其复位

海明窗离散傅里叶变换

$$x_w(n) = x(n) \cdot h(sR - n) \quad (3)$$

以下方程是窗重建序列, 这里 $f(n)$ 代表重建窗

$$x(n) = \frac{1}{M} \sum_{s=-\infty}^{\infty} x_w(n) f(sR-n) \quad (4)$$

2.3 恢复音高改变能量

音高变换处理后,时域窗口的范围也发生了变化。为了准确地回复能量,公式(5)修改了重建窗口。

$$\tilde{x}(n) = \frac{1}{M} \sum_{s=-\infty}^{\infty} x(n) \cdot h_p(sR-n) f_p(sR-n) \quad (5)$$

这里 h_p 音调变系数。 p 是尺度因子窗口。如果用 w_m 来代替 h_p 然后,它可以产生海明重建窗口的系数。 f_p 可以通过比率乘以 f 来构造。这和重建一个重建系数一样。

$$c_p = \frac{M}{\sum_{s=-\infty}^{\infty} w_m(sR-n) w_m[p \cdot (sR-n)]}$$

$$= \frac{M}{\sum_{s=-\infty}^{\infty} (0.54 - 0.46 \cos \frac{2\pi p(sR-n)}{M}) (0.54 - 0.46 \cos \frac{2\pi (sR-n)}{M})}$$

$$= \begin{cases} \frac{R}{(0.08 + 0.92^2 \cdot 1/4)} & p \neq 1 \\ \frac{R}{(0.08 + 0.92^2 \cdot 3/8)} & p = 1 \end{cases} \quad (6)$$

在实际处理中,由于序列的状态可能会被正在处理的改变,综合窗口和堆的处理不能保证恢复所有的能量。但在特定情况下^[9],这个过程可以恢复大部分的能量。实验结果表明,海明窗口比起布莱克曼窗在抑制频率泄漏上更好。在改进算法中,我们使用海明窗口作为分析和综合处理的卷积窗。

2.4 分析/合成处理

我们可以使用一个 FIFO 队列接收音频输入序列。输入队列的长度相当于窗口的长度。输入队列处理在分析每段时间后促进 R 样品。为了抑制频谱走样,它需要窗处理输入队列的顺序。窗口顺序在被处理后将改变成一个新的序列 $\tilde{x}_w(n)$ 。并且新的序列在再次窗计算以恢复原来的信号能量后,将被堆积起来到最后一次移位输出缓冲区。重建的信号是 $\tilde{x}(n)$ 。相声码器^[10]算法使用常数 4 作为重建系数。信号能量在音调改变前后之间不同。使用一个定值不会有一个好的匹配性。改进后的算法引入一个音调因子作为校正因子。它在听感纠正了改变的能量。

2.5 采用快速傅里叶变换进行音频处理

在音频处理中傅立叶变换是最常见的转化形式。傅里叶变换将信号从时域转换到频域,而傅里叶逆变换则将信号从频域转换到时域。在用计算机处理音频时,是无法衡量和计算无限长的信号。正确的方法在傅里叶变换之前是剪取一段时间帧然后运用定期延续的方法去获得一个虚拟的有限信号。截断的信号通过混叠效应传播其能量,这也叫做频率泄漏。在最小化频率

泄露方面用海明窗作为傅里叶变换作为操作元比布莱克曼窗具有更好的性能。在这个过程中,数字音频是模拟信号的离散抽样结果。所以我们一般所说傅里叶变换是指离散傅里叶变换 (DFT),并且其逆方程如下:(7) 傅里叶变换, (8) 逆傅里叶变换。

$$X(k) = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N} \quad (7)$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j2\pi kn/N} \quad (8)$$

$$(k = 0, 1, \dots, N-1)$$

$$mag = \sqrt{real^2 + imag^2}$$

$$phase = \arctan \frac{imag}{real} \quad (9)$$

2.6 频率/相匹配处理

这是改变音调的核心。傅立叶变换后的序列被称为复序列。映射到复平面成为直角坐标。音调变换基于频率改变和相位调制,复杂的序列被表示为模/相极坐标是很有必要的^[7]。

$$mag = \sqrt{real^2 + imag^2}$$

$$phase = \arctan \frac{imag}{real} \quad (10)$$

变换的目的是为了推算光谱成分的两个样品分析序列相差。这两个样本是差异 R 样本。音调变系数乘以相位差产生一个新的相。用这个新相位重组频率复序列^[8]。并映射这个新序列到全面频谱缓冲区。

2.7 主观特征和 M 窗口长度以及跳过样品数 R 三者之间的关系

M 值越大,窗口覆盖的范围越大。并且光谱分析的转换错误越小^[9]。实验表明,人类的耳朵对高音调音乐中的频率错误是很敏感的。一个大窗口可以增加频谱分析的准确性。在 44.1 KHz 的音频中,当 $M > 2048$ 将音高参数调到在 1 到 2 之间可以获得一个较好的效果。当音高系数等于 1 的时候音频音乐保存原来的音高未改变。

2.8 频率调制算法的推广

音频音高变化和时间伸缩(音高不变)可视为一个问题^[10]。一段双采样的音频,当播放音频音高将提高八度,反之亦然。要确保时间改变而音高没有改变,或音高改变而时间没有改变,都必须修改原始音频^[11]。一个简单的例子是,播放已提高八度的音乐采用一半的采样率速度。音高和原来一样,但时间需要两倍。当改变音频的音高时,采用相结合算法然后线性插值或采样,可以得到同样的音高。但是,用新时间范围除以原音频频率可以得到的音频等于原音高变化因子。实验表明,通过相合成它也能达到令人满意的结果^[12]。

3. 总结

2倍波的低频率部分在相位声码器^[4]音高转换算法会被丢失。改进后的算法体现了原波的低频和组成一个新的波形,该波形是原本的两倍。将音乐分为背景声和语音。对于音调上升,如果用频率调制方法将乐器混响直接加到背景音乐则会产生回音效果。对人类的声音可以显然地听到声音的重复;对于音调下降,当用频率调制到两帧之间的连接处时,它会产生的噪音。有效地解决这两个问题的方法是相组成音调变换算法。声音的平滑通过相位声码器可以被减低^[4]。改进算法保持了在主观音乐改变前后之间音乐的响度一致。语音的质量明显地提高了。

因为短期傅里叶变换中固有频率泄的漏影响,合成法法则采用时域窗口减低频谱混叠。直接用频率模型分析帧会导致每个相位不一致。当连接的不同相位,通过用调整相位奇数和总和来处理相变,可以在主观上获得一个美妙的平滑效果。因为窗处理改变了原始序列的能量。但通过复合窗/堆处理可以再次恢复能量。在实际处理中,要根据不同因素选择适当的参数。根据频率上升选择大的窗口长度,根据频率下降选择小的步率。

参考文献:

- [1] J. Laroche, "Time and pitch scale modification of audio signals," in Applications of Digital Signal Processing to Audio and Acoustics, M. Kahrs and K. Brandenburg, Eds. Kluwer, Norwell, MA, 1998.
- [2] J.L. Flanagan and R.M. Golden, "Phase vocoder," Bell Syst. Tech. J., vol. 45, pp. 1493 - 1509, Nov 1966.
- [3] J. B. Allen and L. R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," Proc. IEEE, vol. 65, no. 11,

pp. 1558 - 1564, Nov. 1977.

- [4] R. Portnoff, "Time-scale modifications of speech based on short-time Fourier analysis," IEEE Trans. Acoust., Speech, Signal Processing, vol. 29, no. 3, pp. 374 - 390, 1981.
- [5] M.S. Puckette, "Phase-locked vocoder," in Proc. IEEE ASSP-Workshop on app. of sig. proc. to audio and acous., New Paltz, NY, 1995.
- [6] J. Laroche and M. Dolson, "Improved phase vocoder time-scale modification of audio," to appear in May issue of IEEE trans. speech and audio proc., 1999.
- [7] L.B. Almeida and F.M. Silva, "Variable-frequency synthesis: an improved harmonic coding scheme," in Proc. IEEE Int. Conf. Acoust., Speech, Signal processing, 1984, pp. 27.5.1 - 27.5.4.
- [8] R. J. McAulay and T. F. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, no. 4, pp. 744 - 754, Aug 1986.
- [9] X. Serra and J. Smith, "Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition," Computer Music J., vol. 14, no. 4, pp. 12 - 24, Winter 1990.
- [10] E. B. George and M. J. T. Smith, "Analysis-bysynthesis/Overlap-add sinusoidal modeling applied to the analysis and synthesis of musical tones," J. Audio Eng. Soc., vol. 40, no. 6, pp. 497 - 516, 1992.
- [11] S. Tassart and P. Depalle, "Analytical approximations of fractional delays: Lagrange interpolators and allpass filters," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, Munich, Germany, 1997.
- [12] T.I. Laakso, V. Valimäki, M. Karjalainen, and U. Klaine, "Splitting the unit delay [fir/all pass filters design]," IEEE Signal Processing mag., vol. 13, no.1, pp. 30 - 60, Jan 1996.

(上接第40页)

2.2 仿真结果

对一台实际他励直流电动机进行仿真。该电机的参数^[4]为: $U_N=220V$, $I_N=53A$, $n_N=1500r/min$, $L_a=0.015H$, $R_a=0.5\Omega$, $GD^2=22.45$, $K=0.132$, $C=9.55K=1.261$, $T_L=67N\cdot m$ 。采用全压启动时,电枢电流 I_a 和转速 n 的变化曲线如图3所示;图4是降压启动时电枢电流 I_a 和转速 n 的变化曲线。

由仿真结果可知:与全压启动相比,降压启动的启动电流减小了,另外,两种启动方式对应的稳定转速均能达到额定值,说明所建的直流电动机数学模型是正确的。

3、结束语

本文建立了他励直流电动机的数学模型,并对直流电动机的启动过程进行了仿真研究。仿真结果表明,

将MATLAB软件引入到电机教学过程中,不但提高了学生的学习效率,还充实了该课程的知识结构体系,为深入学习和研究电机提供了一个好的软件平台。

参考文献:

- [1] 王淑芳,刘长青等."电机及驱动技术"课程教材改革方向探讨[J]. 中国电力教育, 2010,1:129-130
- [2] 邓先明,方荣惠等."电机与拖动基础"课程仿真软件的设计[J]. 电气电子教学学报, 2010,2:86-94
- [3] 孔凡才. 自动控制原理与系统 [M]. 北京:机械工业出版社, 2007
- [4] 罗志强. MATLAB在电机仿真中的应用[J]. 电气传动自动化, 2009, 31(4):56-59